



HAL
open science

On the treewidth and related parameters of random geometric graphs

Dieter Mitsche, Guillem Perarnau

► **To cite this version:**

Dieter Mitsche, Guillem Perarnau. On the treewidth and related parameters of random geometric graphs. Symposium on Theoretical Aspects of Computer Science, Feb 2012, Paris, France. hal-00923118

HAL Id: hal-00923118

<https://hal.science/hal-00923118v1>

Submitted on 2 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On treewidth and related parameters of random geometric graphs*

Dieter Mitsche¹ Guillem Perarnau²

1. Dept. of Mathematics, Ryerson University,
350 Victoria St., Toronto, ON, Canada.
dmitsche@ryerson.ca

2. MA4 - Universitat Politècnica de Catalunya,
C/ Jordi Girona 1-3, 08034 Barcelona, Spain.
guillem.perarnau@ma4.upc.edu

January 8, 2013

Abstract

We give asymptotically exact values for the treewidth $\text{tw}(G)$ of a random geometric graph $G \in \mathcal{G}(n, r)$ in $[0, \sqrt{n}]^2$. More precisely, we show that there exists some $c_1 > 0$, such that for any constant $0 < r < c_1$, $\text{tw}(G) = \Theta(\frac{\log n}{\log \log n})$, and also, there exists some $c_2 > c_1$, such that for any $r = r(n) \geq c_2$, $\text{tw}(G) = \Theta(r\sqrt{n})$. Our proofs show that for the corresponding values of r the same asymptotic bounds also hold for the pathwidth and the treedepth of a random geometric graph.

Keywords: Random geometric graphs, treewidth, treedepth, pathwidth.

AMS subject classification: 05C80, 05C62, 90B15.

1 Introduction

Let V be a set of n points in the square $\mathcal{S}_n = [0, \sqrt{n}]^2$ and $r = r(n)$ a nonnegative real number. This choice of the square is only for convenience; by suitable scaling we could have chosen the square $[0, 1]^2$ and all the results would be still valid. We may assume that no two points are placed in the same position, and thus we can identify each point with its position, that is, $v \in V$ refers also to the geometrical position of v in \mathcal{S}_n .

The *geometric graph* G of V with radius r is the graph constructed by connecting two points of V if their euclidean distance in \mathcal{S}_n is smaller than r . For any two points $u, v \in \mathcal{S}_n$ we will denote by $\text{dist}_E(u, v)$ their euclidean distance and by $\text{dist}_G(u, v)$ their distance in the graph G .

Then we define $\mathcal{G}(n, r)$ as the probability space of the geometric graphs of order n with radius r . A graph G chosen uniformly at random from $\mathcal{G}(n, r)$ will be called a *random geometric graph* and will be denoted by $G \in \mathcal{G}(n, r)$. Note that with probability one, no two vertices of $G \in \mathcal{G}(n, r)$ are placed in the same position.

Starting with the seminal paper of Gilbert [7], random geometric graphs have in recent decades received a lot of attention as a model for large communication networks such as sensor

*This research is partially supported by the Catalan Research Council under project 2009SGR1387. The first author is partially supported by the ICT Program of the European Union under contract number 215270 (FRONTS). The second author wants to thank the FPU grant from the Ministerio de Educación de España. An extended abstract of this paper appeared in [16].

networks. Network agents are represented by the vertices of the graph, and direct connectivity is represented by edges. For applications of random geometric graphs, we refer to Chapter 3 of [10], and for a survey of many theoretical results, we refer to Penrose’s monograph [20].

In order to simplify calculations, we will use the well-known idea of Poissonization (see [20, Section 1.7]): we assume that the vertices of $\mathcal{G}(n, r)$ are generated according to a Poisson point process of intensity 1 over \mathcal{S}_n . Conditioned under the fact that this Poisson point process generates exactly n vertices (which happens with probability $\Theta(1/\sqrt{n})$), this model and the $\mathcal{G}(n, r)$ model have the same uniform distribution of the n vertices, and we will use this equivalence from now on. Notice that, using a Poisson point process, the random variables indicating the number of points in disjoint areas of \mathcal{S}_n are independent.

All our stated results are asymptotic as $n \rightarrow \infty$. We use the usual notation *a.a.s.* to denote *asymptotically almost surely*, i.e. with probability $1 - o(1)$. Besides, in order to be safe such that all results from the Poisson model also hold for $\mathcal{G}(n, r)$, all statements that hold *a.a.s.* in fact hold with probability at least $1 - o(1/\sqrt{n})$.

It is well known that the property of the existence of a giant component of order $\Theta(n)$ undergoes a sharp threshold in $\mathcal{G}(n, r)$ (see e.g. [8]), this is, there exists a constant value r_t such that for any $\varepsilon > 0$, *a.a.s.* the largest component of $G \in \mathcal{G}(n, r_t - \varepsilon)$ is of order $O(\log n)$, whereas in $G \in \mathcal{G}(n, r_t + \varepsilon)$, a single component of order $\Theta(n)$ is present, while the others have order $O(\log n)$ (see [20, Chapter 10]). The exact value of r_t is not yet known.

However, there exist two positive constants $c^- \approx \sqrt{0.696}$, $c^+ \approx \sqrt{3.372}$ (see [20], p.189) such that $c^- \leq r_t \leq c^+$. Moreover, simulation studies suggest that the exact value of $r_t \approx \sqrt{1.44}$ (see again [20], p.189).

Since random geometric graphs have been heavily used for modeling communication networks, it is natural to analyze the expected complexity of different algorithms applied to this class. Courcelle’s Theorem [4] states that any problem that can be expressed in monadic second order logic, can be solved in linear time for the class of graphs with bounded treewidth. This motivates the study of this parameter and other tree-like parameters on random geometric graphs. In this paper, we study the behavior of the treewidth and the treedepth on random geometric graphs.

The treewidth was introduced independently by Halin in [9] and by Robertson and Seymour in [24].

For a graph $G = (V, E)$ on n vertices, we call (T, \mathcal{W}) a *tree decomposition* of G , where \mathcal{W} is a set of vertex subsets $W_1, \dots, W_s \subseteq V$, called bags, and T is a forest with vertices in \mathcal{W} , such that

1. $\bigcup_{i=1}^s W_i = V$.
2. For any $e = uv \in E$ there exists a set $W_i \in \mathcal{W}$ such that $u, v \in W_i$.
3. For any $v \in V$, the subgraph induced by the $W_i \ni v$ is connected as a subgraph of T .

The *width* of a tree-decomposition is $w(T, \mathcal{W}) = \max_{1 \leq i \leq s} |W_i| - 1$, and the *treewidth* of a graph G can be defined as

$$\text{tw}(G) = \min_{(T, \mathcal{W})} w(T, \mathcal{W}).$$

Observe that if G is a graph with connected components H_1, \dots, H_m , then

$$\text{tw}(G) = \max_{1 \leq i \leq m} \text{tw}(H_i). \tag{1}$$

The concept of treedepth has been introduced under different names in the literature. In this paper we follow the definition given by Nešetřil and Ossona de Mendez as a tree-like parameter in the scope of homomorphism theory, where it provides an alternative definition of bounded expansion classes [18]. For the sake of completeness, we note that the treedepth is

also equivalent to the height of an elimination tree (used for instance in the parallel Cholesky decomposition [22]). Furthermore, analogous definitions can be found using the terminology of rank function [17], vertex ranking number (or ordered coloring) [6] or weak coloring number [11].

We now give the precise definition of treedepth. Let T be a rooted tree. The *height* of T is defined as the number of vertices of the longest rooted path. The *closure* of T is the graph that has the same set of vertices and a pair of vertices is connected by an edge if one is an ancestor of the other in T . We say that the tree T is an *elimination tree* of a connected graph G if G is a subgraph of its closure. The *treedepth* of a connected graph G , $\text{td}(G)$, is defined to be the minimum height of an elimination tree of G .

The definition of treedepth can also be extended to nonconnected graphs. If G is a graph with connected components H_1, \dots, H_m ,

$$\text{td}(G) = \max_{1 \leq i \leq m} \text{td}(H_i). \quad (2)$$

Hence, if $S \subset V(G)$ separates G in two subsets A and B , we have

$$\text{td}(G) \leq |S| + \max\{\text{td}(A), \text{td}(B)\}. \quad (3)$$

Observe that if H is a subgraph of G , then

$$\text{td}(H) \leq \text{td}(G) \text{ and } \text{tw}(H) \leq \text{tw}(G). \quad (4)$$

Both parameters are closely connected: while the treewidth of a graph G is a parameter that measures the similarity between G and the class of trees in general, the treedepth of G measures how close G is to a star. In other words, the treedepth also takes into account the diameter of the tree we are comparing the graph with. The two parameters are related by the following inequalities:

$$\text{tw}(G) \leq \text{td}(G) \leq (\text{tw}(G) + 1) \log_2 n,$$

both bounds being sharp (see [18]). Note also that $\text{tw}(G) \geq \omega(G) - 1$, where $\omega(G)$ denotes the size of the largest clique in G .

Results of the paper. In this paper we study the values of $\text{tw}(G)$ and $\text{td}(G)$ of a random geometric graph $G \in \mathcal{G}(n, r)$ for different values of $r = r(n)$. In particular, we prove the following two main theorems:

Theorem 1. *There is some positive constant $c_1 < c^-$, such that for any $0 < r \leq c_1$ and $G \in \mathcal{G}(n, r)$, a.a.s. $\text{tw}(G) = \Theta(\frac{\log n}{\log \log n})$, and also a.a.s. $\text{td}(G) = \Theta(\frac{\log n}{\log \log n})$.*

Theorem 2. *There is some constant $c_2 > c^+$, such that for any $r = r(n) \geq c_2$ and $G \in \mathcal{G}(n, r)$, a.a.s. $\text{tw}(G) = \Theta(r\sqrt{n})$, and also a.a.s. $\text{td}(G) = \Theta(r\sqrt{n})$.*

Remark 1. Inspecting the proof of Theorem 1, one can easily see that, adapting constants, Theorem 1 holds up to c^- . In the case of Theorem 2, however, this is not clear to us.

Remark 2. For $G \in \mathcal{G}(n, r)$ with r constant, but $r \geq c_2$, by the results of [5], many problems such as STEINER TREE, FEEDBACK VERTEX SET, CONNECTED VERTEX COVER can be solved in time $O(\text{poly}(n)3^{\sqrt{n}})$, while others like CONNECTED DOMINATING SET, CONNECTED FEEDBACK VERTEX SET, MIN CYCLE COVER, LONGEST PATH, LONGEST CYCLE, GRAPH METRIC TRAVELLING SALESMAN PROBLEM can be solved in time $O(\text{poly}(n)4^{\sqrt{n}})$.

Remark 3. Other width parameters that are sandwiched between the treewidth and the treedepth clearly then also have the same asymptotic behavior in $\mathcal{G}(n, r)$. For instance, the *pathwidth* of a graph, introduced by Robertson and Seymour [23], measures the similarity between a graph and a path. Since the pathwidth is well known to be bounded from below by the treewidth and bounded from above by the treedepth (see Theorem 5.3 and Theorem 5.11 of [25]), the former theorems imply that for those values of $r = r(n)$ the pathwidth of the graph is of the same order.

Remark 4. Whereas intuitively it might be clear that around the threshold of the existence of a giant component there should be a jump for parameters like treewidth or treedepth in $\mathcal{G}(n, r)$, the orders of magnitude of these parameters are not so obvious (for us). Moreover, we point out that there are differences between $\mathcal{G}(n, r)$ and $\mathcal{G}(n, p)$: it is known that in the Erdős-Rényi random graph model $\mathcal{G}(n, p)$, as soon as the giant component appears, the graph has linear treewidth (see [14]). In contrast to this, Theorem 2 shows that a random geometric graph with a linear number of edges containing a giant component only has treewidth $\Theta(\sqrt{n})$. This different behavior of the two models can be explained by their different expansion properties and the connection between balanced separators and treewidth (see Lemma 14 below). given by Lemma 14. Classical random graphs have very good expansion properties, and thus it is difficult to find small separators of large sets of vertices. The geometric properties of the model $\mathcal{G}(n, r)$ imply the lack of large expanders. For this reason, in the latter case one can construct a tree decomposition with smaller bags. On the other hand, in the subcritical regime (with a linear number of edges, but before the existence of a giant component) the treedepth of $\mathcal{G}(n, p)$ is $\Theta(\log \log n)$ (see [21]), whereas by Theorem 1, for random geometric graphs it is already $\Theta(\frac{\log n}{\log \log n})$. Furthermore, in this range, in classical random graphs the treewidth is bounded by a constant (see [21]), whereas our theorems show that in $\mathcal{G}(n, r)$ both treewidth and treedepth are asymptotically of the same order for a wide range of parameters r . The fact that for random geometric graphs the treedepth and treewidth are always asymptotically equal implies that $\mathcal{G}(n, r)$ is more similar to a star-shaped tree than to a path-shaped tree, which in general is not true for random graphs.

The paper is organized as follows. In Section 2 we define the cell graph of a geometric graph and give some properties of it. The proof of Theorem 1 is presented in Section 3. Whereas the lower bound follows from a standard argument using the clique number of $\mathcal{G}(n, r)$, the proof of the upper bound is more involved. In Section 4 we continue by proving Theorem 2. Finally, in Section 5 we conclude mentioning some open problems.

2 Properties of Deterministic Geometric Graphs

2.1 The cell graph of a geometric graph

For any constant $\ell > 0$, we can tessellate \mathcal{S}_n into squares of length ℓ called *cells*. We use this tessellation to construct the *cell graph* $C_G(\ell)$ of G : each nonempty cell will be represented by a *vertex* and two different vertices of $C_G(\ell)$ will be joined if there exist two points of G in the corresponding cells that share an edge (see Figure 1, where the tessellation is omitted for clarity).

For the sake of simplicity of the presentation, we assume that \sqrt{n}/ℓ is an integer for the values of ℓ considered in this paper. From now on, unless otherwise stated, we will call *points* to the vertices of the geometric graph G and use the word *vertex* for the cells of $C_G(\ell)$. The cell-graph $C_G(\ell)$ simplifies the original geometric graph G while preserving the same structure. For any subgraph H of G we will denote its cell graph by $C_H(\ell)$.

Remark 5. Notice that $C_H(\ell)$ is always a subgraph of $C_G(\ell)$. Observe that, for any $\ell \leq r/\sqrt{2}$, each nonempty cell contains points from exactly one connected component of G , since all the points inside a cell are connected. Thus, if $\ell \leq r/\sqrt{2}$ there exists a natural bijection between the connected components of G and the connected components of $C_G(\ell)$.

We need another auxiliary graph, the *grid graph* $L_{a,b}^k$, defined as follows; its vertex set is $V(L_{a,b}^k) = \{(i, j) : 1 \leq i \leq a, 1 \leq j \leq b\}$, and $(i, j)(i', j') \in E(L_{a,b}^k)$ if and only if $(i, j) \neq (i', j')$ and $\max\{|i - i'|, |j - j'|\} \leq k$. Note that by construction, for a geometric graph G in \mathcal{S}_n with radius r we have

$$C_G(\ell) \subseteq L_{\sqrt{n}/\ell, \sqrt{n}/\ell}^{\lceil r/\ell \rceil}, \quad (5)$$

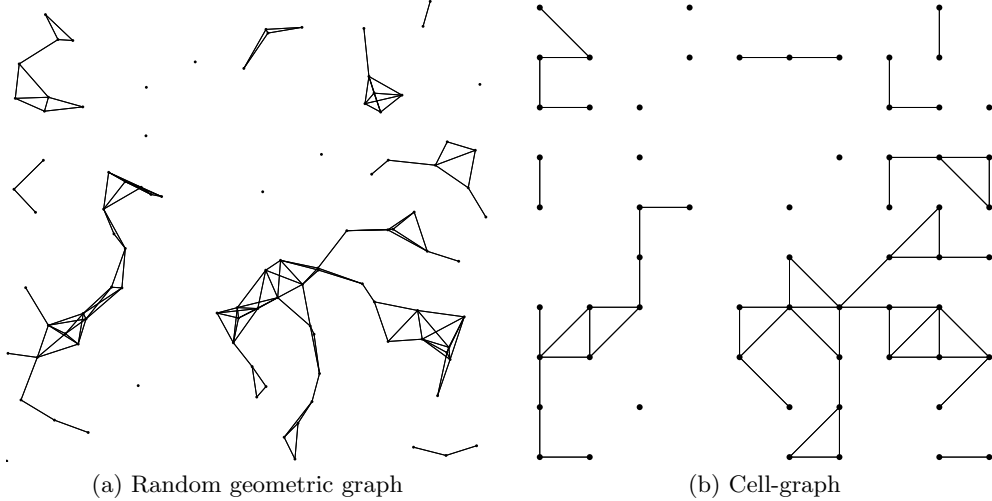


Figure 1: A random geometric graph and its corresponding cell graph

contained as a subgraph.

The following lemma bounds the maximal number of different connected subgraphs of a given size in $L_{a,b}^k$.

Lemma 3. *The number of connected subgraphs of size s in $L_{a,b}^k$ is at most $O(ab(2k+1)^{4s})$.*

Proof. A connected subgraph is determined by a root v and any of its spanning trees, rooted at v . Observe that there are ab many ways to choose $v \in V(L_{a,b}^k)$. Moreover, the degree of a vertex in $L_{a,b}^k$ is at most $(2k+1)^2$, since for any cell (i, j) there are at most $(2k+1)^2$ cells (i', j') such that $\max\{|i - i'|, |j - j'|\} \leq k$.

One can construct at most $((2k+1)^2)^{2s-3} \leq (2k+1)^{4s}$ walks of length $2s-2$ that have both start and end points at v . In particular, these walks contain all the possible spanning trees rooted at v since a spanning tree has $s-1$ edges and each edge is traversed twice. Thus, the lemma follows. \square

Remark 6. Lemma 3 is certainly not tight. For the same problem on the integer lattice (each cell is connected to the four closest ones) the asymptotic growth is $\text{poly}(s)\lambda^s$. However the exact value of λ is not yet known. The best known lower and upper bounds for λ are 3.980137 and 4.65, respectively (see [2, 12]).

The following proposition bounds the treedepth of a strong product of a graph and a clique. Given two graphs G_1 and G_2 , the strong product $G = G_1 \boxtimes G_2$ is defined as $V(G) = V(G_1) \times V(G_2)$ and $(u_1, u_2)(v_1, v_2) \in E(G)$ iff for $i = 1, 2$, either $u_i = v_i$ or $u_i v_i \in E(G_i)$. Denote by K_t the complete graph on t vertices.

Lemma 4. *Let $G = G_1 \boxtimes K_t$. Then*

$$\text{td}(G) \leq t \text{td}(G_1).$$

Proof. Let T_1 be a tree of height $\text{td}(G_1)$ that embeds G_1 in its closure. Note also that K_t is contained in the closure of a rooted path of order t , P_t . Observe that $T_1 \boxtimes P_t$ is not a tree, but it contains a tree T , in whose closure $T_1 \boxtimes P_t$ is contained (see Figure 2). Indeed, T can be constructed in the following way: each vertex $u \in V(T_1)$ is replaced by a path of order t (call these new vertices u_1, \dots, u_t), and if there is an edge $uv \in E(T_1)$, such that u is ancestor of v , then in T , u_t is connected by an edge to v_1 (the depth of v_1 in T is exactly one more than the depth of u_t), see Figure 2. Note that T is a tree and its closure contains G as a subgraph. Since each vertex of G_1 is replaced by t vertices, $\text{td}(G) \leq t \text{td}(G_1)$. \square

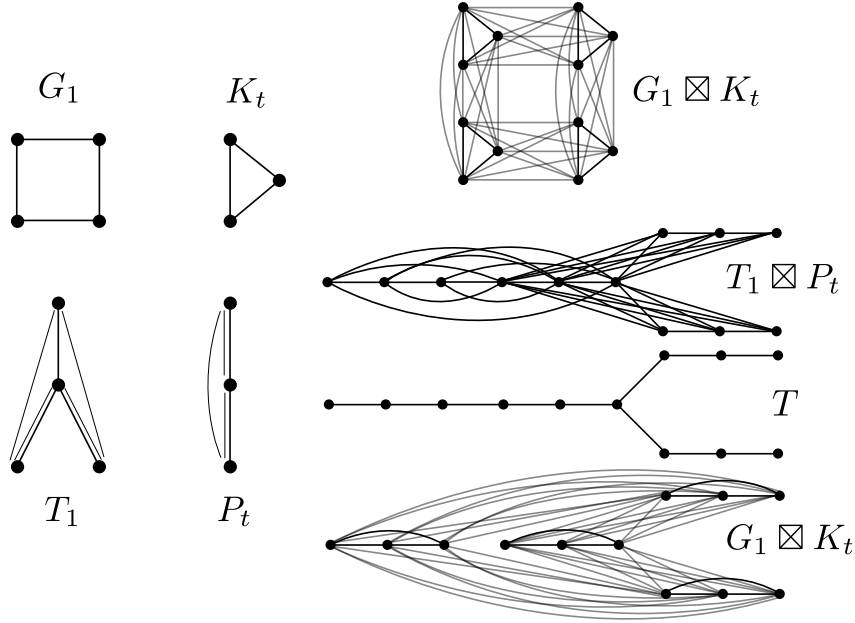


Figure 2: Embedding of the strong product.

Observe also that for a geometric graph G ,

$$G \subseteq C_G(\ell) \boxtimes K_t, \quad (6)$$

where t is the maximum number of points inside a cell of the tessellation of length ℓ .

Since we can express the treedepth of G in terms of the treedepth of its cell graph and the latter one is a subgraph of $L_{a,b}^k$, the following proposition will be useful,

Proposition 5. *Let $L_{a,b}^k$ the grid graph defined as above and suppose that $a \leq b$. Then*

$$\text{td}(L_{a,b}^k) \leq O(ka \log b).$$

Proof. We present an elimination tree for $L_{a,b}^k$ in a recursive way. First, note that $\text{td}(L_{a,k}^k) = O(ka)$, since the treedepth of a graph is always smaller than its order. Let us compute now the treedepth of $L_{a,b}^k$. By removing the central copy of $L_{a,k}^k$ in $L_{a,b}^k$, we disconnect the original graph and we get two copies of $L_{a,(b-k)/2}^k$. Applying this recursively and using (3), we obtain

$$\text{td}(L_{a,b}^k) \leq O(ka) + \text{td}(L_{a,(b-k)/2}^k) \leq \dots \leq \underbrace{O(ka) + \dots + O(ka)}_{\log b} + \text{td}(L_{a,k}^k) = O(ka \log b).$$

□

The following proposition will be very useful in the proof of Theorem 1, but can be applied to any sparse geometric graph.

Proposition 6. *Let H be a geometric graph of order m such that there are no more than t points inside each cell of length $\ell = r/\sqrt{2}$.*

Then, we have

$$\text{td}(H) = O\left(\max\left\{\frac{m}{\log m}, t(\log m)^3\right\}\right).$$

Proof. Throughout this proof all the cells will have length $\ell = r/\sqrt{2}$. Notice that by Remark 5 the connected components of the cell graph $C_H(\ell)$ are in one to one correspondence with the connected components in H . Thus, we may assume that H is connected. We will show an upper bound on $\text{td}(H)$ by providing an elimination scheme for C_H which then induces an elimination scheme for H .

Fix a vertex $v \in V(C_H)$ corresponding to a cell of the tessellation. For any integer $d \geq 0$, denote by V_d the set of vertices in the cell graph at distance d from v . More precisely,

$$V_d = \{u \in V(C_H) : \text{dist}_{C_H}(v, u) = d\} .$$

Analogously, we define P_d to be the set of points of H inside the cells of V_d .

For the sake of convenience, we define

$$K = \frac{m}{(\log m)^2} .$$

The idea of the proof is to find a separator S of H that contains at most $O(K)$ points. This separator will split the graph into some smaller subgraphs. Using (3) and applying the same procedure recursively to the remaining parts, we will get an upper bound on $\text{td}(H)$.

Let f be the largest integer for which

$$\sum_{d=0}^{f-1} |P_d| \leq \frac{m}{2} . \quad (7)$$

Let f_1 be the largest integer for which $f_1 \leq f$ and $|P_{f_1}| \leq K$ and f_2 be the smallest integer for which $f_2 \geq f$ and $|P_{f_2}| \leq K$. Since H contains m points, $f_2 - f_1 \leq \frac{m}{K} = (\log m)^2$.

Given a graph G and $S \subset V(G)$, we will denote by $G[S]$ the subgraph of G induced by S . We decompose C_H into the following subgraphs (see Figure 3):

$$C_S = C_H[V_{f_1} \cup V_{f_2}] , \quad C_A = C_H \left[\bigcup_{d=0}^{f_1-1} V_d \right] , \quad C_L = C_H \left[\bigcup_{d=f_1+1}^{f_2-1} V_d \right] \quad \text{and} \quad C_B = C_H \left[\bigcup_{d \geq f_2+1} V_d \right] ,$$

and we define accordingly

$$H_S = H[P_{f_1} \cup P_{f_2}] , \quad H_A = H \left[\bigcup_{d=0}^{f_1-1} P_d \right] , \quad H_L = H \left[\bigcup_{d=f_1+1}^{f_2-1} P_d \right] \quad \text{and} \quad H_B = H \left[\bigcup_{d \geq f_2+1} P_d \right] .$$

Note also that in the case $|P_f| \leq K$, we have $f_1 = f_2$ and C_L and H_L are graph on zero vertices. Suppose this is not the case, and focus on C_L .

Since $\ell = r/\sqrt{2}$, by (5) we know that C_L is a subgraph of at most 4 copies of $L_{a,b}^2$ (see Figure 3), where $a = (\log m)^2$ and $b = m$, since $f_2 - f_1 \leq (\log m)^2$ and $|P_d| \leq m$ for any d . By (3) and Proposition 5, we get

$$\text{td}(C_L) \leq O(4a) + \text{td}(L_{a,b}^2) = O((\log m)^3) .$$

Moreover, $H_L \subseteq C_L \boxtimes K_t$. Hence, by Lemma 4,

$$\text{td}(H_L) = O(t(\log m)^3) .$$

By (3), now applied to H and the separator $S = V_{f_1} \cup V_{f_2}$, we have

$$\begin{aligned} \text{td}(H) &\leq |S| + \max\{\text{td}(H_A), \text{td}(H_L), \text{td}(H_B)\} \\ &\leq 2K + \max\{\text{td}(H_A), O(t(\log m)^3), \text{td}(H_B)\}, \end{aligned} \quad (8)$$

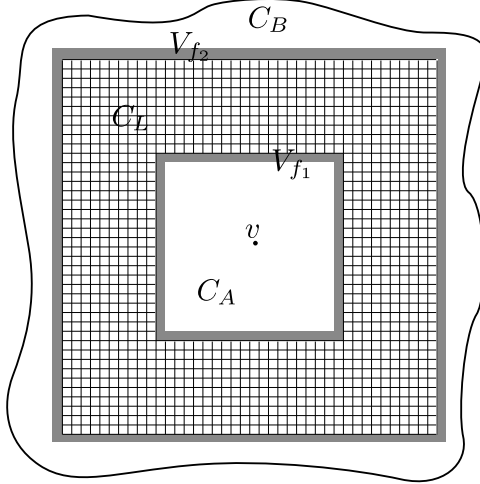


Figure 3: Decomposition of C_H

since $|S| \leq 2K$ by definition of f_1 and f_2 .

We recursively repeat this procedure for the two subgraphs H_A and H_B . By the choice of f in (7), both subgraphs contain at most $m/2$ points. Hence, the recursion depth of our procedure is at most $\log_2 m = O(\log m)$. This implies that

$$\text{td}(H) = O\left(\max\{K \log m, t(\log m)^3\}\right) = O\left(\max\left\{\frac{m}{\log m}, t(\log m)^3\right\}\right).$$

□

2.2 Separators and cells

During the rest of the section we will consider a tessellation of length $\ell = r/4$.

Given $\mathcal{S} \subseteq \mathcal{S}_n$ a set with positive measure, we denote by $\text{vol}(\mathcal{S})$ the area of \mathcal{S} and by $\partial\mathcal{S}$ its boundary in the euclidean topology. We also use $\text{vol}(\partial\mathcal{S})$ to refer to the length of $\partial\mathcal{S}$.

For any set $A \subseteq V(H)$, let $\mathcal{A} = \{x \in \mathcal{S}_n : \min_{v \in A} \text{dist}_E(x, v) \leq \frac{r}{2}\} \subseteq \mathcal{S}_n$, and notice that $\partial\mathcal{A} = \{x \in \mathcal{S}_n : \min_{v \in A} \text{dist}_E(x, v) = \frac{r}{2}\}$.

We will use the fact that for any cell D and for any two elements $u, v \in D$

$$\text{dist}_E(u, v) \leq \frac{r}{2\sqrt{2}}. \quad (9)$$

The following lemma shows that for any separator S of a geometric graph H , we can find a large number of cells of length $\ell = r/4$, whose points are entirely contained in S (see also Figure 4, left).

Lemma 7. *Let H be a connected geometric graph of order m and $S \subset V(H)$ be a separator of H . Fix a connected component H_1 of $H \setminus S$ and denote by $A = V(H_1)$.*

If $\text{vol}(\mathcal{A}) < cn$ for some $c < 1$, there exists a set of cells D_S of size d_S , such that all points inside D_S belong to S and

$$d_S = \Omega\left(r^{-1} \sqrt{\text{vol}(\mathcal{A})}\right).$$

Proof. Define $B = V(H) \setminus (S \cup A)$, the set of vertices that are contained neither in S nor in A .

For any pair of points $v \in A$ and $w \in B$, we have that $\text{dist}_E(v, w) \geq r$, since they are in different connected components of $H \setminus S$. Denote the boundary of \mathcal{A} by $\mathcal{C} = \partial\mathcal{A}$ and observe that since all the points in \mathcal{C} lie at distance exactly $r/2$ from some point at A , they lie at distance at least $r/2$ from any point in B .

Let D_S be the union of the cells that have nonempty intersection with \mathcal{C} . Let us point out that some of these cells may not contain any point of $V(H)$. By Lemma 8 shown below, $d_S = \Omega(r^{-1}\sqrt{\text{vol}(\mathcal{A})})$.

Moreover, all the points contained in D_S belong to S : by (9), any point u contained in D_S lies at distance at most $r/(2\sqrt{2})$ from some element $c \in \mathcal{C}$. However, all the points of $A \cup B$ lie at distance at least $r/2$ from all the elements of \mathcal{C} . Thus, $u \notin A \cup B$, implying that $u \in S$. \square

Remark 7. Suppose that the separator set S induces t connected components in H . Since $\ell = r/4$, by (5), C_H is contained in the grid graph $L_{4\sqrt{n}/r, 4\sqrt{n}/r}^4$. Consider the subgraph of this grid graph, induced by the cells in D_S . Observe that this subgraph forms at most t connected components, at most one for each connected component of S . Denote these connected components by D_{C_j} and their respective sizes by d_{C_j} . Then, by concavity of the square root function,

$$d_S = \sum_j d_{C_j} = \Omega\left(r^{-1}\sqrt{\text{vol}(\mathcal{A})}\right).$$

It remains to prove Lemma 8.

Lemma 8. *With the notation of Lemma 7, we have*

$$d_S = \Omega\left(r^{-1}\sqrt{\text{vol}(\mathcal{A})}\right).$$

Proof. Denote again by $\mathcal{C} = \partial\mathcal{A}$. We make use of the following isoperimetric inequality (see [19], Theorem 1.6.1): for any set of positive measure $\mathcal{S} \subset \mathbb{R}^2$,

$$\text{vol}(\partial\mathcal{S}) \geq \Omega(\sqrt{\text{vol}(\mathcal{S})}). \quad (10)$$

The inequality in [19] is stated for connected sets \mathcal{S} , but by concavity of the square root function it holds for general \mathcal{S} . In particular, if $\text{vol}(\mathcal{S}) \leq cn$ for some constant $c < 1$, the same inequality holds for $\mathcal{S} \subset \mathcal{S}_n$. By hypothesis of Lemma 7, $\text{vol}(\mathcal{A}) < cn$ for some constant $c < 1$, thus,

$$\text{vol}(\mathcal{C}) = \text{vol}(\partial\mathcal{A}) \geq \Omega(\sqrt{\text{vol}(\mathcal{A})}). \quad (11)$$

Recall that D_S is defined to be the set of cells that have nonempty intersection with \mathcal{C} . For any cell $D \in D_S$ we denote by $\mathcal{C}_D = \mathcal{C} \cap D$, the restriction of \mathcal{C} to D . We will show that the length of \mathcal{C}_D is not too large by projecting the elements of \mathcal{C}_D onto ∂D , in such a way that the length of the curve does not decrease by too much.

Let $p : \mathcal{C}_D \rightarrow \partial D$ the application that sends an element $c \in \mathcal{C}_D \subset \mathcal{C}$ being at distance $r/2$ from a point $v \in A$ to the intersection of ∂D and the segment that joins c and v (see Figure 4, right). In case that there is more than one point of A at the same distance from c , $p(c)$ chooses one of them arbitrarily.

Note that p is injective, since no two elements of \mathcal{C}_D can have the same image: indeed, suppose that there exist two different $c, c' \in \mathcal{C}_D$ with corresponding points $v, v' \in A$ such that $p(c) = p(c')$. Then, the segments cv and $c'v'$ would cross at $p(c)$, and either $\text{dist}_E(c, v') < r/2$ or $\text{dist}_E(c', v) < r/2$ holds, contradicting the definition of \mathcal{C} .

Let us show that the application is not too much contracting. Recall that $\text{dist}_E(c, v) = r/2$. Since $c, p(c) \in D$, by (9) we have $\text{dist}_E(c, p(c)) \leq \frac{r}{2\sqrt{2}}$, and therefore $\text{dist}_E(p(c), v) \geq \frac{\sqrt{2}-1}{2\sqrt{2}}r$ by the triangle inequality. A simple geometric argument shows that

$$\text{vol}(p(\mathcal{S})) \geq \frac{\frac{r}{2}}{\frac{\sqrt{2}-1}{2\sqrt{2}}\pi r} \text{vol}(\mathcal{S}).$$

Since p is injective and $\text{vol}(\partial D) = r$,

$$\text{vol}(\mathcal{C}_D) = O(\text{vol}(\partial D)) = O(r) .$$

Using this upper bound for all the cells $D \in D_S$,

$$d_S \geq \frac{\text{vol}(\mathcal{C})}{\max_{D \in D_S} \text{vol}(\mathcal{C}_D)} = \Omega \left(r^{-1} \sqrt{\text{vol}(\mathcal{A})} \right) .$$

□

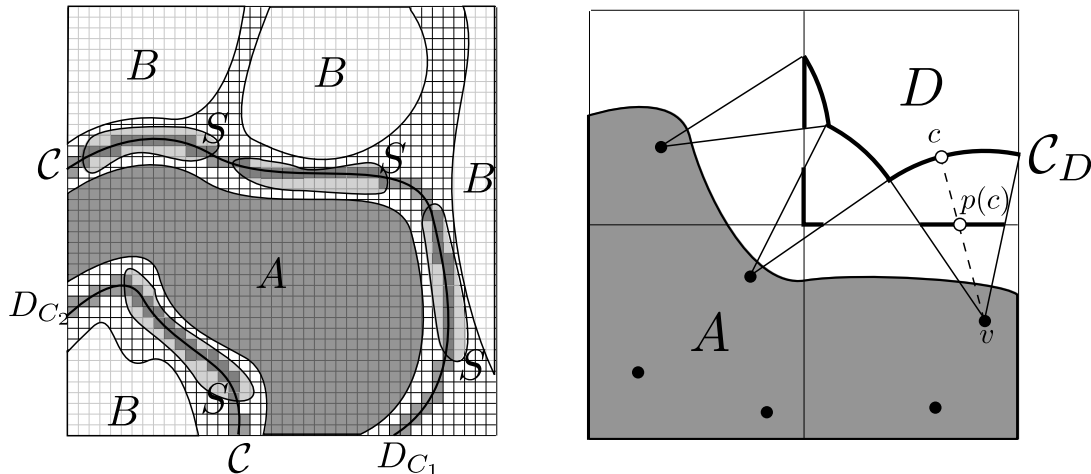


Figure 4: Cells of D_S and the projection of \mathcal{C}_D .

We finish with some properties of the tessellation with $\ell = r/4$.

Lemma 9. *Let H be a geometric graph with connected components H_1, \dots, H_t . Define $A_i = V(H_i)$ and consider a tessellation with $\ell = r/4$. Then, for any cell D*

1. *if there exists a point $v \in A_i$ such that $v \in D$, $D \subset A_i$.*
2. *there are at most 24 curves $\mathcal{C}_i = \partial A_i$ that intersect the cell.*

Proof. For the first part, by (9), for any $u \in D$,

$$\text{dist}_E(u, v) < \frac{r}{2} ,$$

and thus $u \in A_i$.

For the second part, observe that if \mathcal{C}_i intersects D , then there must exist a point of $v \in A_i$ at distance at most $r/2$ from some point in D . There are most 24 cells satisfying this criterion, namely the ones in the first and second neighborhood of D . Since all the points of a cell belong to the same component (they are all connected), there are at most 24 different curves \mathcal{C}_i intersecting D . □

3 Subcritical regime

In this section we compute the treedepth of a random geometric graph with $r < c_1$, that is, below the existence of a giant component. The constant c_1 will be chosen in such a way that the order of each component is *a.a.s.* at most $\log n$ (this value exists, see Theorem 10.3 of [20], and

is only chosen to simplify calculations). We also assume $r = \Theta(1)$. We will use the following result several times: McDiarmid in [15] proved that for any $r = \Theta(1)$ and $G \in \mathcal{G}(n, r)$, *a.a.s.*

$$\omega(G) = (1 + o(1)) \frac{\log n}{\log \log n}, \quad (12)$$

In fact, by looking at the results closely, both results can be seen to hold with probability at least $1 - o(n^{-1/2})$.

By (2), the order of the largest connected component implies a coarse upper bound, namely

$$\text{td}(G) = O(\log n).$$

In order to find a better upper bound, more work is needed. First, we need the following simple lemma, whose proof is included for completeness.

Lemma 10. *Let X be a random variable that follows a Poisson distribution with parameter λ . Then, for any $k \geq 2\lambda$,*

$$\Pr(X \geq k) \leq 2\Pr(X = k).$$

Proof.

$$\begin{aligned} \Pr(X \geq k) &= \sum_{i \geq k} \Pr(X = i) = \sum_{i \geq k} e^{-\lambda} \frac{\lambda^i}{i!} \\ &= e^{-\lambda} \frac{\lambda^k}{k!} \left(1 + \frac{\lambda}{k+1} + \frac{\lambda^2}{(k+1)(k+2)} + \dots \right) \\ &\leq e^{-\lambda} \frac{\lambda^k}{k!} \sum_{i \geq 0} \left(\frac{\lambda}{k} \right)^i = e^{-\lambda} \frac{\lambda^k}{k!} \frac{1}{1 - \frac{\lambda}{k}} \\ &\leq 2e^{-\lambda} \frac{\lambda^k}{k!} = 2\Pr(X = k), \end{aligned}$$

where the last inequality follows from the assumption $k \geq 2\lambda$. □

For the sake of convenience, we define

$$T_{\max} = \frac{2 \log n}{\log \log n} \quad \text{and} \quad T = \frac{\sqrt{2 \log n}}{\log \log n}.$$

From now on, we consider in this section the cell graph $C_G(\ell)$ of $G \in \mathcal{G}(n, r)$ with $\ell = r/\sqrt{2}$ and write simply C_G for $C_G(\ell)$. Notice that all the points inside a cell of C_G form a clique. Hence, by (12), each cell contains less than T_{\max} points *a.a.s.* For this particular tessellation, we call a cell *sparse* if it contains less than T points, and *dense* otherwise.

Proposition 11. *Every connected component H of $G \in \mathcal{G}(n, r)$ contains at most $O(T_{\max})$ points in dense cells with probability at least $1 - o(n^{-1/2})$.*

Proof. For any connected component H of G we will show that the probability that the number of points in dense cells of H is at least $10T_{\max}$ is $o(n^{-3/2})$. Since there are clearly at most n connected components in G , by taking a union bound over all them, with probability $1 - o(n^{-1/2})$ no component will have more than $10T_{\max}$ points in dense cells.

Let A_i be the number of points in the cell i . Since we are using a Poisson point process of intensity 1, A_i follows a Poisson distribution with parameter $\lambda = r/2$, the expected number of points in that cell. The probability that a cell is dense can be expressed as $p = \Pr(A_i \geq T)$.

By Lemma 10,

$$(1 - O(T^{-1})) \frac{e^{-\lambda}}{\sqrt{2\pi T}} \left(\frac{e\lambda}{T} \right)^T = \Pr(A_i = T) \leq p = \Pr(A_i \geq T) \leq 2\Pr(A_i = T) \leq \frac{2e^{-\lambda}}{\sqrt{2\pi T}} \left(\frac{e\lambda}{T} \right)^T, \quad (13)$$

where we have used Stirling approximation $T! = (1 + O(T^{-1}))\sqrt{2\pi T} \left(\frac{T}{e}\right)^T$.

To count the number of points lying in dense cells, we define the following random variables for each cell $i \in V(C_G)$:

$$Y_i = \begin{cases} t & \text{if } i \text{ is dense and has } t \text{ points inside,} \\ 0 & \text{otherwise.} \end{cases}$$

Our aim is to show that $Y_H = \sum_{i \in V(C_H)} Y_i$ is at most $O(T_{\max})$.

Notice that the probability that the cell i is sparse is $1 - p$, while the probability of having $T + j$ points is

$$\Pr(A_i = T + j) = (1 - O((T + j)^{-1})) \frac{e^{-\lambda}}{\sqrt{2\pi(T+j)}} \left(\frac{e\lambda}{T+j}\right)^{T+j} \leq \left(\frac{e\lambda}{T}\right)^T \frac{e^{-\lambda}}{\sqrt{2\pi T}} \left(\frac{e\lambda}{T}\right)^j,$$

for any integer $j \geq 0$. Using (13) we have

$$\Pr(A_i = T + j) \leq 2p \left(\frac{e\lambda}{T}\right)^j.$$

These observations lead to the definition of the following independent random variables (R_i):

$$R_i = \begin{cases} 0 & \text{with probability } 1 - 2p, \\ T + j & \text{with probability } 2p \left(\frac{e\lambda}{T}\right)^j \text{ for any } j \geq 1, \\ T & \text{with probability } 2p \left(1 - \frac{e\lambda}{T-e\lambda}\right). \end{cases}$$

First of all, observe that R_i is a probability distribution. The random variables Y_i and R_i have similar distributions. In particular, each variable R_i stochastically dominates the corresponding random variable Y_i . Analogously, we define $R = \sum_{i \in V(C_H)} R_i$. Then,

$$\Pr(R \geq j) \geq \Pr(Y \geq j), \quad (14)$$

for any $j \geq 0$. In particular, this also holds, if $j = O(T_{\max})$.

Therefore, it is enough to compute an explicit upper bound for $\Pr(R \geq 10T_{\max})$. By the assumption on c_1 , H contains at most $\log n$ points, and thus, there are $|V(C_H)| \leq \log n$ vertices in C_H . Since we aim for an upper bound, we may assume that $|V(C_H)| = \log n$.

Since $\ell = r/\sqrt{2}$, by (5) we know that $C_H(\ell)$ is a subgraph of $L^2_{\sqrt{2n}/r, \sqrt{2n}/r}$. By Lemma 3 applied with $s = \log n$ there are at most $O(n5^{4\log n}) < n^8$ ways to construct C_H .

By a union bound over all the possible combinations that yield $R > 10T_{\max}$, we have

$$\Pr(R > 10T_{\max}) \leq n^8 \sum_{m=1}^{\log n} \sum_{S \in \binom{V(C_H)}{m}} \sum_{\sum_{i \in S} c_i \geq 10T_{\max}} \Pr\left(\bigcap_{i \in S} R_i = c_i\right), \quad (15)$$

where m counts the number of dense cells in the distribution given by the R_i , S is the set of dense cells and c_i is the number of points inside the dense cell $i \in S$. There are at most $(\log n)^m$ ways to choose the set S and at most $(T_{\max})^m < (\log n)^m$ possible values for c_i .

Recall that the variables R_i are independent and that $\Pr(R_i = T + j) = 2p \left(\frac{e\lambda}{T}\right)^j$ for any $j \geq 1$. Therefore,

$$\Pr\left(\bigcap_{i \in S} R_i = c_i\right) = \prod_{i=1}^m 2p \left(\frac{e\lambda}{T}\right)^{c_i - T}.$$

On the one hand, if $m \leq 10\sqrt{\log n}$, using (13),

$$\prod_{i=1}^m 2p \left(\frac{e\lambda}{T}\right)^{c_i - T} \leq \prod_{i=1}^m \frac{4}{\sqrt{2\pi T}} \left(\frac{e\lambda}{T}\right)^{c_i} \leq \prod_{i=1}^m \left(\frac{e\lambda}{T}\right)^{c_i} \leq (2e^\lambda \sqrt{2\pi T} p)^{\sum c_i} \leq (2e^\lambda \sqrt{2\pi T} p)^{10\sqrt{\log n}}. \quad (16)$$

On the other hand, if $m = 10\sqrt{\log n} + j$ for some integer $j \geq 1$,

$$\prod_{i=1}^m 2p \left(\frac{e\lambda}{T} \right)^{c_i - T} \leq (2p)^m = (2p)^{10\sqrt{\log n}} (2p)^j.$$

Therefore, splitting equation (15) into two sums, we obtain

$$\begin{aligned} \Pr(R > 10T_{\max}) &\leq n^8 \sum_{m=1}^{10\sqrt{\log n}} (\log n)^{2m} (2e^\lambda \sqrt{2\pi T} p)^{10\sqrt{\log n}} \\ &\quad + n^8 (2(\log n)^2 p)^{10\sqrt{\log n}} \sum_{j \geq 1} (2(\log n)^2 p)^j. \end{aligned}$$

From the bounds on p in (13), one can derive that $(2(\log n)^2 p) < 1/2$, and the infinite sum of the second term above is bounded from above by one. Thus,

$$\begin{aligned} \Pr(R > 10T_{\max}) &\leq n^8 \left(10\sqrt{\log n} \right) \left((\log n)^2 p (2e^\lambda \sqrt{2\pi T} + 2) \right)^{10\sqrt{\log n}} \\ &= \exp \left\{ 8 \log n + 5 \log \log n + 10\sqrt{\log n} (2 \log \log n + \log p + O(\log T)) \right\}. \end{aligned}$$

Moreover, by (13) we also have $p \leq \frac{2e^{-\lambda}}{\sqrt{2\pi T}} \left(\frac{e\lambda}{T} \right)^T$, and hence $\log p \leq -(1 + o(1))T \log T \leq -\sqrt{\log n}$. Thus,

$$\Pr(R > 10T_{\max}) < \exp \{ -(1 + o(1))2 \log n \} = o(n^{-3/2}). \quad (17)$$

By (14), this also implies that $\Pr(Y > 10T_{\max}) = o(n^{-3/2})$, and by taking a union bound over all components, this implies that the probability of having a connected component with more than $10T_{\max}$ points inside dense cells is $o(n^{-1/2})$. \square

Proof of Theorem 1. The lower bound on $\text{tw}(G)$ follows easily from (12), which yields

$$\text{td}(G) \geq \text{tw}(G) \geq \omega(G) - 1 = \Omega \left(\frac{\log n}{\log \log n} \right).$$

For the upper bound, we provide a constructive way to create an elimination tree for G . By (2) it suffices to bound from above the treedepth of each connected component. Let H be a connected component of G .

From Proposition 11, there are at most $O(T_{\max})$ points in dense cells of H . We temporarily remove all these points, and add them at the end. Let H' be the subgraph of H that remains after removing the points in the dense cells.

Observe that now, by definition of sparse, every cell of $C_{H'}$ contains at most T points. Denoting by $m = |V(H')|$, by Proposition 6 we have

$$\text{td}(H') = O \left(\max \left\{ \frac{m}{\log m}, T(\log m)^3 \right\} \right),$$

Since, with probability at least $1 - o(n^{-3/2})$, $m = O(\log n)$, we have that for every component H of G , $\text{td}(H') = O(T_{\max})$ with probability at least $1 - o(n^{-1/2})$.

Recall that adding a new point to H can increase the treedepth by at most one unit. Thus, $\text{td}(H) \leq \text{td}(H') + O(T_{\max}) = O(T_{\max})$, and therefore, using (1), we have

$$\text{td}(G) = O \left(\frac{\log n}{\log \log n} \right)$$

with probability at least $1 - o(n^{-1/2})$. \square

4 Supercritical regime

Fix now $r = r(n) \geq c_2$, for some sufficiently large constant $c_2 \geq c^+ \geq r_t$, the threshold radius of having a giant component.

Recall that for any subset $\mathcal{S} \subseteq \mathcal{S}_n = [0, \sqrt{n}]^2$ with positive measure, we denote by $\text{vol}(\mathcal{S})$ the area of A . We need the following standard lemma (which is a simple application of Chernoff bounds for Poisson variables, see for example Theorem A.1.15 of [1]):

Lemma 12. *For any $\mathcal{S} \subseteq \mathcal{S}_n$ such that $\text{vol}(\mathcal{S}) \geq c \log n$ and any $0 < \delta < 1/3$, the number of points inside \mathcal{S} is*

1. *at most $(1 + \delta) \text{vol}(\mathcal{S})$ with probability at least $1 - (e^\delta (1 + \delta)^{-(1+\delta)})^{\text{vol}(\mathcal{S})} \geq 1 - e^{-\frac{\delta^2}{3} \text{vol}(\mathcal{S})}$.*
2. *at least $(1 - \delta) \text{vol}(\mathcal{S})$ with probability at least $1 - e^{-\frac{\delta^2}{2} \text{vol}(\mathcal{S})}$.*

We will use the previous lemma to show that there exist separating sets with few points, and consequently, give an upper bound on $\text{td}(G)$.

Proposition 13. *For any $r \geq c_2$, $\text{td}(G) \leq O(r\sqrt{n})$ with probability $1 - e^{-\Omega(r\sqrt{n})}$.*

Proof. Consider the tessellation of \mathcal{S}_n into square cells of length $\ell = r$. Denote by $D_{(i,j)}$ the j -th cell in the i -th row, where $1 \leq i, j \leq a = \sqrt{n}/r$.

Define

$$X_1^1 = \left(\bigcup_{i=1}^a D_{(a/2,i)} \right) \cup \left(\bigcup_{i=1}^a D_{(i,a/2)} \right),$$

and consider the set $Y_1^1 \subset V(G)$, containing the points inside X_1^1 . Observe that Y_1^1 is a separator, since $\ell = r$, and it splits the graph into 4 components (some of them might be empty), G_2^1 , G_2^2 , G_2^3 and G_2^4 .

By Equation (3), we have

$$\text{td}(G) \leq |Y_1^1| + \max_{1 \leq j \leq 4} \{\text{td}(G_2^j)\}.$$

We then define analogously the sets X_2^j , for all G_2^j , and using Equation (3), we continue iteratively. Let t denote the step where all the sets X_t^j have size one (see Figure 5).

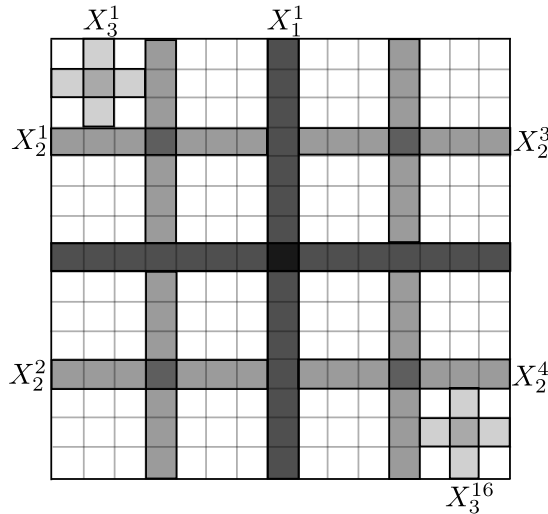


Figure 5: Construction of the sets X_i^j .

The treedepth of G will be bounded from above by the maximum number of points inside any of the possible sets of cells

$$X_{j_1 j_2 \dots j_t} = X_1^{j_1} \cup X_2^{j_2} \cup \dots \cup X_t^{j_t}$$

where $1 \leq j_i \leq 4^{i-1}$.

Observe that $|X_i^{j_i}| \leq a2^{-(i-2)}$. The sets $X_{j_1 j_2 \dots j_t} = X_1^{j_1} \cup X_2^{j_2} \cup \dots \cup X_t^{j_t}$ are not disjoint, but they all have the same size

$$|X_{j_1 j_2 \dots j_t}| = \sum_{i=1}^t |X_i^{j_i}| \leq \sum_{i=1}^t a2^{-(i-2)} = (1 - 2^{-t})4a.$$

Let $Y_{j_1 j_2 \dots j_t}$ denote the set of points in $X_{j_1 j_2 \dots j_t}$. Thus, $|Y_{j_1 j_2 \dots j_t}|$ is a random variable following a Poisson distribution with mean at most $(1 - 2^{-t})4ar^2$.

By Lemma 12.1,

$$\begin{aligned} \Pr(|Y_{j_1 j_2 \dots j_t}| \geq (1 + \delta)(1 - 2^{-t})4ar^2) &< e^{-(1-2^{-t})4\delta^2 ar^2/3} \\ &= e^{-\Omega(r\sqrt{n})}, \end{aligned}$$

for any $0 < \delta < 1/3$.

Moreover, there are at most

$$\prod_{i=1}^t 4^{i-1} = e^{O(t^2)}$$

sets of the form $X_{j_1 j_2 \dots j_t}$. Observe also that, by construction, $t = O(\log a) = O(\log n)$.

Now, by a union bound over all sets,

$$\Pr(\exists j_1, j_2, \dots, j_t : |Y_{j_1 j_2 \dots j_t}| > (1 + \delta)(1 - 2^{-t})4ar^2) \leq e^{O(\log^2 n) - \Omega(r\sqrt{n})} = e^{-\Omega(r\sqrt{n})}.$$

Thus, we have that the treedepth of G is at most

$$\text{td}(G) \leq (1 + \delta)(1 - 2^{-t})4ar^2 = O(r\sqrt{n}),$$

with probability at least $1 - e^{-\Omega(r\sqrt{n})}$. □

For a lower bound on $\text{tw}(G)$, we need the following explicit link between the treewidth of a graph and the existence of a vertex separator with special properties. A vertex partition $V = (A, S, B)$ is a *balanced k -partition* if $|S| = k + 1$, S separates A and B , and $\frac{1}{3}(n - k - 1) \leq |A|, |B| \leq \frac{2}{3}(n - k - 1)$. In this case, S is also called a *balanced separator*. The following result connecting balanced partitions and treewidth is due to Kloks [13].

Lemma 14 ([13]). *Let G be a graph with n vertices and $\text{tw}(G) \leq k$ such that $n \geq k - 4$. Then G has a balanced k -partition.*

From now on and until the end of the section, we consider the tessellation of \mathcal{S}_n into square cells of size $r/4$.

Recall that for any set $A \subset V(H)$, we defined $\mathcal{A} = \{x \in \mathcal{S}_n : \min_{v \in A} \text{dist}_E(x, v) \leq r/2\}$. Observe that in a geometric graph, no direct relation exists between the size of A and the volume of \mathcal{A} . However, as the following lemma shows, in a random geometric graph, $\text{vol}(\mathcal{A})$ can be bounded from below using the size of A , if A has linear size.

Lemma 15. *Let $G \in \mathcal{G}(n, r)$ with $r \geq c_2$. For any set $A \subseteq V(G)$, $|A| \geq \alpha n$, there exists $c(\alpha) > 0$, such that*

$$\text{vol}(\mathcal{A}) \geq c(\alpha)n,$$

with probability $1 - e^{-\Omega(n)}$.

Proof. Set $m = m(\alpha)$ to be the smallest constant integer such that

$$\frac{e^{-1}}{m!} \left(\frac{m^2}{m-1} + \frac{m}{(m-1)^2} \right) \leq \frac{\alpha}{8} \quad \text{and } m \geq 4e,$$

which exists for any $\alpha > 0$, since the left-hand side of the first condition tends to zero, when $m \rightarrow +\infty$.

Recall that the number of points inside a cell D follows a Poisson distribution with mean $\lambda = r^2/16$. Suppose that D contains $t \geq 0$ points. Define then Z_D to be the following random variable:

$$Z_D = \begin{cases} t & \text{if } t \geq m\lambda, \\ 0 & \text{otherwise;} \end{cases}$$

and let $Z = \sum Z_D$ be the sum of these random variables over all the cells of the tessellation.

We may consider $r \geq 4$, since by hypothesis $r \geq c_2$, for some c_2 large enough. This implies that $\lambda \geq 1$. By Stirling bounds and calculation of the derivative one can see that for any $m \geq 1$ the function $f(\lambda) = e^{-\lambda} \frac{\lambda^{m\lambda}}{(m\lambda)!}$ is decreasing for $\lambda \in [1, \infty)$, and thus

$$\Pr(Z_D = m\lambda) = e^{-\lambda} \frac{\lambda^{m\lambda}}{(m\lambda)!} \leq \frac{e^{-1}}{m!}.$$

Also

$$\Pr(Z_D = m\lambda + i) = e^{-\lambda} \frac{\lambda^{m\lambda+i}}{(m\lambda+i)!} = e^{-\lambda} \frac{\lambda^{m\lambda+(i-1)}}{(m\lambda+(i-1))!} \cdot \frac{\lambda}{m\lambda+i} \leq \frac{1}{m} \Pr(Z_D = m\lambda + (i-1)),$$

for any $i \geq 1$. Hence,

$$\mathbb{E}(Z_D) = \sum_{t \geq m\lambda} t \Pr(Z_D = t) \leq \frac{e^{-1}}{m!} \sum_{i \geq 0} (m\lambda + i) m^{-i} \leq \frac{e^{-1}}{m!} \left(\frac{m^2\lambda}{m-1} + \frac{m}{(m-1)^2} \right) \leq \frac{\alpha\lambda}{4},$$

where the last inequality follows from the definition of m . Since $\lambda = r^2/16$ and there are $16n/r^2$ cells in the tessellation, we have

$$\mathbb{E}(Z) \leq \frac{\alpha n}{4}.$$

By Hoeffding bounds for unbounded random variables (the precise version we use here is Theorem 1 of [3], applied with $X_D = \epsilon_D = Z_D$, and thus $S = T = Z$, $Y = Po(\lambda)$, $m_k = m = \mathbb{E}(Z_D)$ for any k , and $b = m\lambda - 1$, so that $m(b) = m$ and the measure $\mu^{[m]}$ is exactly our probability distribution of Z_D , and $x = 2\mathbb{E}(Z)$)

$$\Pr(Z > 2\mathbb{E}(Z)) < \inf_{h < x} e^{-h2\mathbb{E}(Z)} \mathbb{E}(e^{hZ}) \leq e^{-2\mathbb{E}(Z)} \mathbb{E}(e^Z).$$

Now, observe that

$$e^{2\mathbb{E}(Z_D)} \geq e^{2m\lambda} \Pr(Z_D = m\lambda) \geq e^{(2m-1)\lambda} \frac{\lambda^{m\lambda}}{(m\lambda)!}$$

and

$$\mathbb{E}(e^{Z_D}) = \Pr(Z_D = 0) + \sum_{i \geq 0} e^{m\lambda+i} \Pr(Z_D = m\lambda + i) \leq 1 + e^{(m-1)\lambda} \frac{\lambda^{m\lambda}}{(m\lambda)!} \sum_{i \geq 0} \left(\frac{e}{m} \right)^i.$$

Since by assumption on m , $e/m \leq 1/4$, we have

$$\mathbb{E}(e^{Z_D}) \leq 1 + \frac{4}{3} \frac{\lambda^{m\lambda}}{(m\lambda)!} e^{(m-1)\lambda} \leq \frac{3}{2} \frac{\lambda^{m\lambda}}{(m\lambda)!} e^{(m-1)\lambda}.$$

The random variables Z_D are mutually independent. Thus,

$$e^{2\mathbb{E}(Z)} = \prod e^{2\mathbb{E}(Z_D)} \geq \left(\frac{\lambda^{m\lambda}}{(m\lambda)!} e^{(2m-1)\lambda} \right)^{\frac{16n}{r^2}}$$

and

$$\mathbb{E}(e^Z) \leq \left(\frac{3}{2} \frac{\lambda^{m\lambda}}{(m\lambda)!} e^{(m-1)\lambda} \right)^{\frac{16n}{r^2}},$$

and therefore

$$\Pr(Z > 2\mathbb{E}(Z)) \leq e^{-2\mathbb{E}(Z)} \mathbb{E}(e^Z) \leq \left(\frac{3}{2} e^{-m\lambda} \right)^{\frac{16n}{r^2}} = e^{-\Omega(n)}.$$

Thus, with probability at least $1 - e^{-\Omega(n)}$, there are at most $\alpha n/2$ points of G contained in cells with at least $m\lambda$ points, and thus with the same probability there are at least $\alpha n/2$ points of A contained in cells with less than $m\lambda$ points.

Therefore, with this probability, there are at least

$$\frac{\alpha n/2}{m\lambda} = \frac{8\alpha n}{mr^2},$$

different cells D that contain at least one point from A . By Lemma 9.1, $D \subset \mathcal{A}$, and

$$\text{vol}(\mathcal{A}) \geq \frac{8\alpha n}{mr^2} \cdot \text{vol}(D) = c(\alpha)n,$$

with probability at least $1 - e^{-\Omega(n)}$. □

Using the previous lemmata, we are able to provide a lower bound for $\text{tw}(G)$.

Theorem 16. *There exists a constant c_2 such that for any $r \geq c_2$, $\text{tw}(G) \geq \Omega(r\sqrt{n})$ with probability at least $1 - e^{-\Omega(r\sqrt{n})}$.*

Proof. We will show that there exists no balanced separator of size $o(r\sqrt{n})$ for the giant component H . Then, by Lemma 14, this implies that $\text{tw}(H) = \Omega(r\sqrt{n})$, and by Equation (1), $\text{tw}(G) \geq \text{tw}(H) = \Omega(r\sqrt{n})$.

Let $S \subset V(H)$ be a fixed balanced separator of H . Let S_1, \dots, S_t denote subsets of S that induce connected components in H . We may assume that S is minimal, and hence each component of S contains at least one point of H . Therefore we can assume that $t < r\sqrt{n}$, as otherwise there is nothing to prove.

The separator S is balanced and $|S| = o(r\sqrt{n}) = o(n)$. Thus, there exist two sets $A, B \subset V(H)$ (not necessarily connected) of size $\frac{1}{3}n(1 - o(1)) \leq |A|, |B| \leq \frac{2}{3}n(1 + o(1))$, such that $H \setminus S$ contains no edges from A to B .

Since both sets have linear size, we may assume that $|A| \geq \alpha n$ and $|B| \geq \beta n$. Thus, by Lemma 15, with probability at least $1 - e^{-\Omega(n)}$, their respective volumes are also linear, which implies that

$$c(\alpha)n \leq \text{vol}(\mathcal{A}) \leq (1 - c(\beta))n,$$

with at least this probability, where $c(\alpha), c(\beta) > 0$ are given by Lemma 15.

Since the complementary event has probability only $e^{-\Omega(n)}$ and we will below take union bounds over a set of at most $e^{vr\sqrt{n}}$ events, we condition on this event from now on.

Noticing that $\text{vol}(\mathcal{A}) \leq (1 - c(\beta))n$, we can apply Lemma 7 to the separator S and each connected component of A separately. Then, by concavity of the square root function, we have a set of cells D_S of size

$$d_S = \Omega\left(r^{-1} \sqrt{\text{vol}(\mathcal{A})}\right) \geq \frac{K\sqrt{n}}{r},$$

for some constant $K > 0$, such that all the points inside D_S belong to S .

Now it suffices to show that, with high probability, there are at least $r\sqrt{n}$ points inside any potential set of cells D_S . Denote by Y_{D_S} the random variable counting the number of points inside such a D_S . Since $\text{vol}(D_S) = \frac{r^2}{16}d_S$, by Lemma 12.2, we have that

$$\Pr\left(Y_{D_S} < (1 - \delta)\frac{r^2}{16}d_S\right) \leq e^{-\frac{\delta^2 r^2}{32}d_S}. \quad (18)$$

We will now show that by taking a union bound over all possible balanced separators no balanced separator contains less than $(1 - \delta)\frac{r^2}{16}d_S$ points. Write $D_S = \cup D_{C_j}$, where D_{C_j} are the sets of cells described in Remark 7. Recall that there are at most t such sets and we will assume that this is the case. For the sake of simplicity, we will assume that $r \geq 4$. By letting d_{C_1}, \dots, d_{C_t} the sizes of these connected components of D_S , and setting $a = b = 4\sqrt{n}/r \leq \sqrt{n}$, $k = 4$ and $s = d_{C_j}$ in Lemma 3, we conclude that there are at most $n^t 9^{4(d_{C_1} + \dots + d_{C_t})} \leq n^t e^{9d_S}$ ways to construct D_S .

Combining the inequality in (18) with a union bound over all separators D_S of size $d_S \geq K\sqrt{n}/r$, the probability of having such a balanced separator is at most

$$\Pr(\exists S : S \text{ is balanced sep., } |S| = o(r\sqrt{n})) \leq \sum_{d_S \geq K\sqrt{n}/r} \sum_{t \leq r\sqrt{n}} \sum_{d_{C_1} + \dots + d_{C_t} = d_S} n^t e^{9d_S} e^{-\gamma r^2 d_S}, \quad (19)$$

where $\gamma = \delta^2/32$ for any $0 < \delta < 1/3$.

Our aim for the rest of the proof is to show that each summand can be bounded from above by an exponentially small term.

The number of ways to obtain the sum d_S when using t nonnegative numbers is at most $d_S^t \leq n^t$, and thus, the right hand side of (19) can be bounded from above by

$$\sum_{d_S \geq \sqrt{n}/r} \sum_{t \leq r\sqrt{n}} n^{2t} e^{9d_S} e^{-\gamma r^2 d_S}. \quad (20)$$

If $t \leq c \frac{r\sqrt{n}}{\log n}$ for some small constant $c > 0$, we can bound $n^{2t} < e^{2cr\sqrt{n}} = o(e^{\gamma r^2 d_S})$, and also $e^{9d_S} = o(e^{\gamma r^2 d_S})$, for sufficiently large r and not too small γ .

Thus, we can assume that $t > c \frac{r\sqrt{n}}{\log n}$. Denote by D_H all cells that contain at least one point of H . Suppose first that there is a constant fraction of the cells in $D_H \setminus D_S$ contained in components of size at least $\frac{\sqrt{n} \log n}{cr}$. We restrict our separator to these big components. For this (sub)separator, since there are at most $16n/r^2$ cells, we have $t \leq c \left(\frac{\sqrt{n}}{r \log n}\right)$, and by the previous arguments, for this (sub)separator, the probability of having too few points is at most $e^{-\gamma r^2 d_S}$ for some $\gamma > 0$, and hence the probability of having few points in S is also at most $e^{-\gamma r^2 d_S}$.

Thus, we may assume that there is at least a constant fraction of the cells in $D_H \setminus D_S$ contained in components of order at most $\frac{\sqrt{n} \log n}{cr}$. Then, d_S is minimized if there are at most $c' \frac{\sqrt{n}}{r \log n}$ components of order $\frac{\sqrt{n} \log n}{cr}$.

Observe that, by Lemma 9.2, there exist at most 24 different connected components A_i of A , such that ∂A_i intersects a given cell. Hence, by applying the isoperimetric inequality given in (10) over each component,

$$d_S \geq \frac{n^{1/4} \sqrt{\log n}}{\sqrt{cr}} \times \frac{c'}{24 r \log n} \sqrt{n} = \Omega\left(\frac{n^{3/4}}{r^{3/2} \sqrt{\log n}}\right).$$

We distinguish two cases.

First, we consider the case $c_2 \leq r = O(\sqrt{\log n})$. Since $t \leq r\sqrt{n}$, $n^{2t} = e^{2t \log n} \leq e^{2r\sqrt{n} \log n} \leq e^{2\sqrt{n} \log^{3/2} n}$ and $e^{\gamma r^2 d_S} \geq e^{\gamma \frac{n^{3/4} \sqrt{r}}{\sqrt{\log n}}} \geq e^{\gamma \frac{n^{3/4}}{\sqrt{\log n}}}$,

$$n^{2t} e^{9d_S} e^{-\gamma r^2 d_S} \leq e^{-\gamma' r^2 d_S}$$

for some $0 < \gamma' < \gamma$.

Otherwise, $r = \omega(\sqrt{\log n})$. Observe that $t \leq d_S$, since $d_{C_j} \geq 1$ by definition. Therefore,

$$n^{2t} e^{9d_S} e^{-\gamma r^2 d_S} \leq n^{2d_S} e^{9d_S} e^{-\gamma r^2 d_S} = e^{(2 \log n + O(1) - \gamma r^2) d_S} \leq e^{-\gamma'' r^2 d_S}$$

for some $0 < \gamma'' < \gamma$.

We showed that each term of (20) can be bounded by an exponentially small term. Hence, there exists a constant $\nu > 0$, such that with probability at most

$$\sum_{d_S \geq K\sqrt{n}/r} \sum_{t \leq r\sqrt{n}} n^{2t} e^{9d_S} e^{-\nu r^2 d_S} = e^{-\Omega(r\sqrt{n})},$$

there exists a separator S containing less than $(1 - \delta) \frac{r^2}{16} d_S = \Omega(r\sqrt{n})$ points connected to the giant component, completing the proof. \square

Proof of Theorem 2. Theorem 2 follows directly by recalling that $\text{tw}(G) \leq \text{td}(G)$ and combining Proposition 13 with Theorem 16. \square

5 Conclusion

We have shown that for random geometric graphs with $0 < r \leq c_1$ and for $r \geq c_2$ the parameters of treewidth and treedepth are asymptotically of the same order. The immediate natural question that remains open is whether for all values of $r = \Theta(1)$, including the values of $c_1 \leq r \leq c_2$, this happens to be true. For either of the parameters it would be interesting to know whether there is a sharp threshold width of order $o(1)$, in the sense that there exists some critical value of the radius r_c such that the treewidth (treedepth, respectively) of a graph with radius of at most $r_c - o(1)$ is of order $\Theta(\frac{\log n}{\log \log n})$ with probability at least $1 - \epsilon$, and the treewidth (treedepth, respectively) of a graph with radius at least $r_c + o(1)$ is of order $\Theta(\sqrt{n})$ with probability at least $1 - \epsilon$, for any $\epsilon > 0$. We remark that the general result on sharp thresholds of monotone properties of [8] implies only a sharp threshold width of order $\log^{3/4} n$. Needless to say, in case of the existence of such a sharp threshold, it would be nice to find this exact threshold value for any of the two parameters (they might coincide). Using our methods, this, however, among other problems, requires the knowledge of the exact threshold value r_t of the appearance of the giant component in a random geometric graph, which at the moment is not known.

References

- [1] Noga Alon and Joel Spencer, *The probabilistic method*, John Wiley and Sons, 2008.
- [2] Gill Barequet, Micha Moffie, Ares Ribó, and Günter Rote, *Counting polyominoes on twisted cylinders*, *Integers* **6** (2006), A22, 37. MR 2247816 (2008g:05046)
- [3] Vidmantas Bentkus, *An extension of the hoeffding inequality to unbounded random variables*, *Lith. Math. J.* **48** (2008), no. 2, 137–157.
- [4] Bruno Courcelle, *The monadic second-order logic of graphs. I. Recognizable sets of finite graphs*, *Inform. and Comput.* **85** (1990), no. 1, 12–75.

- [5] Marek Cygan, Jesper Nederlof, Marcin Pilipczuk, Michal Pilipczuk, Johan van Rooij, and Jakub Onufry Woitaszczyk, *Solving connectivity problems parameterized by treewidth in single exponential time*, ArXiv e-prints (2011).
- [6] Jitender S. Deogun, Ton Kloks, Dieter Kratsch, and Haiko Müller, *On vertex ranking for permutation and other graphs*, STACS 94 (Caen, 1994), Lecture Notes in Comput. Sci., vol. 775, Springer, Berlin, 1994, pp. 747–758.
- [7] Edward N. Gilbert, *Random plane networks*, Journal of the Society for Industrial and Applied Mathematics **9** (1961), 533–543.
- [8] Ashish Goel, Sanatan Rai, and Bhaskar Krishnamachari, *Sharp thresholds for monotone properties in random geometric graphs*, STOC, 2004, pp. 580–586.
- [9] Rudolf Halin, *S-functions for graphs*, Journal of Geometry **8** (1976), 171–186 (English).
- [10] Ramin Hekmat, *Ad-hoc networks-fundamental properties and network topologies*, Springer, 2006.
- [11] Hal A. Kierstead and Daqing Yang, *Orderings on graphs and game coloring number*, Order **20** (2003), no. 3, 255–264 (2004).
- [12] David A. Klarner and Ronald L. Rivest, *A procedure for improving the upper bound for the number of n -ominoes*, Tech. report, Stanford, CA, USA, 1972.
- [13] Ton Kloks, *Treewidth*, Lecture Notes in Computer Science, vol. 842, Springer-Verlag, Berlin, 1994, Computations and approximations.
- [14] Choongbum Lee, Joonkyung Lee, and Sang-il Oum, *Rank-width of random graphs*, Journal of Graph Theory **70** (2012), no. 3, 339–347.
- [15] Colin McDiarmid, *Random channel assignment in the plane*, Random Structures Algorithms **22** (2003), no. 2, 187–212. MR 1954610 (2003m:94004)
- [16] Dieter Mitsche and Guillem Perarnau, *On the treewidth and related parameters of random geometric graphs*, STACS, 2012, pp. 408–419.
- [17] Jaroslav Nešetřil and Saharon Shelah, *On the order of countable graphs*, European J. Combin. **24** (2003), no. 6, 649–663.
- [18] Jaroslav Nešetřil and Patrice Ossona de Mendez, *Tree-depth, subgraph coloring and homomorphism bounds*, European J. Combin. **27** (2006), no. 6, 1022–1041.
- [19] John Oprea, *Differential geometry and its applications*, (2007).
- [20] Mathew Penrose, *Random geometric graphs*, Oxford Studies in Probability, vol. 5, Oxford University Press, Oxford, 2003.
- [21] Guillem Perarnau and Oriol Serra, *On the tree-depth of Random Graphs*, accepted to Discrete Applied Mathematics (2012).
- [22] Alex Pothen and Sivan Toledo, *Handbook on data structures and applications*, ch. 59, Elimination structures in scientific computing, Chapman and Hall, 2004.
- [23] Neil Robertson and Paul D. Seymour, *Graph minors. I. Excluding a forest*, Journal of Combinatorial Theory **35** (1983), no. 1, 39–61.
- [24] ———, *Graph minors. II. Algorithmic aspects of tree-width*, J. Algorithms **7** (1986), no. 3, 309–322.

[25] Robert Sasak, *Comparing 17 graph parameters*, Master Thesis, Univ. of Bergen, 2010.